BIVARIATE NATURAL EXPONENTIAL FAMILIES WITH QUADRATIC DIAGONAL OF THE VARIANCE FUNCTION

JOANNA MATYSIAK

Abstract. We characterize bivariate natural exponential families having the diagonal of the variance function of the form

$$\mathrm{diag}V(m_1, m_2) = \left(Am_1^2 + am_1 + bm_2 + e, Am_2^2 + cm_1 + dm_2 + f\right),$$

with $A < 0$ and $a, \ldots, f \in \mathbb{R}$. The solution of the problem relies on finding the conditions under which a specific parametric family of functions consists of Laplace transforms of some probability measures.

1. Introduction

Let us recall some basic concepts concerning natural exponential families (NEFs for short). For a positive measure μ on $\mathbb{R}^n$ we define its Laplace transform by

$$L_\mu(\theta) = \int_{\mathbb{R}^n} \exp\langle\theta, x\rangle\, \mu(dx).$$

Let $\Theta(\mu) = \mathrm{Int}\{\theta \in \mathbb{R}^n : L_\mu(\theta) < \infty\}$ and let $k_\mu = \log L_\mu$ denote the cumulant function of μ. Set $\mathcal{M}$ as the set of such measures μ that $\Theta(\mu) \neq \emptyset$ and μ is not concentrated on an affine hyperplane of $\mathbb{R}^n$. For $\mu \in \mathcal{M}$, the family of probabilities

$$F(\mu) = \{P(\theta, \mu)(dx) = \exp\left(< \theta, x > -k_\mu(\theta)\right)\mu(dx) : \theta \in \Theta(\mu)\}$$

is called *the natural exponential family generated by* μ, see [15]. The image M_F of $\Theta(\mu)$ by diffeomorphism k'_μ is called *the domain of means of* μ. *The variance function* (VF) of the NEF is defined by $V_F(\mathbf{m}) = k''_\mu(\Psi_\mu(\mathbf{m}))$, where $\mathbf{m} \in M_F$ and $\Psi_\mu = (k'_\mu)^{-1}$. In what follows we will be frequently omitting the subscript F in the notation V_F for the variance function.

The principal significance of the mapping $\mathbf{m} \mapsto V_F(\mathbf{m})$ is that, together with its domain M_F, it characterizes NEF uniquely. It makes it possible to describe NEFs assuming a concrete form of the variance function. There has been a lot of interest in such questions. For example, Letac [16] characterizes all NEFs with $V_F(\mathbf{m}) = B\mathbf{m} + C$, where B is a linear operator mapping $\mathbb{R}^n$ into $\mathbb{S}_n$ (the space of $n \times n$ symmetric real matrices), and $C \in \mathbb{S}_n$. Casalis [4], [5] gives a generalization of this result by considering $V_F(\mathbf{m}) =$

2000 *Mathematics Subject Classification.* Primary: 44A10; Secondary: 60E10.
Key words and phrases. Laplace transform, probability measures, variance functions, regression conditions for probability measures.

$a\mathbf{m}\otimes\mathbf{m}+B\mathbf{m}+C$. Her result was further extended by Hassairi and Zarai [8] to NEFs with cubic variance functions.

For another standpoint, we recall some papers that provide characterizations based on a fragmentary knowledge of the variance function. Kokendji and Seshadri [10] start with $\det V(\mathbf{m}) = const$ and thus identify the Gaussian law in $\mathbb{R}^n$. In [14] Letac and Wesołowski characterized NEFs with $V(\mathbf{m}) = p^{-1}\mathbf{m}\otimes\mathbf{m} - \phi(\mathbf{m})M_\nu$, where M_ν is a symmetric matrix associated with a quadratic form ν, $\mathbf{m}\to\phi(\mathbf{m})$ an unknown real function, and p is a number.

One of the most important papers that base on a partial knowledge of V_F is [1], in which the authors considered the diagonal family of NEFs in $\mathbb{R}^n$ such that

$$\mathrm{diag}V(\mathbf{m}) = (f_1(m_1),\ldots,f_n(m_n))$$

(diag$V$ stands for the diagonal elements of the variance matrix V, and f_i is an arbitrary function of i-th coordinate of $\mathbf{m}$). They gave the full characterization of the family (showing also that f_i, $i=1,2,\ldots,n$ must be some polynomials of degree at most 2).

A natural question is the following: let us fix n classes of functions $\mathcal{F}_i$ of the variables $m=(m_1,m_2,\ldots,m_n)$ for $i=1,\ldots,n$. For instance $\mathcal{F}_i$ could be the functions of the form $f_i(m) = Am_i^2 + B(m) + c$, where B is a linear form and $A, c\in\mathbb{R}$. What happens if we stay focused on the diagonal $(f_1(m),\ldots,f_n(m))$ of the VF such that $f_i\in\mathcal{F}_i$? We are going to study this question for $n=2$. For a partial result in this direction, see [6].

The paper is organized as follows. Its main results, Theorem 2.2 and Theorem 2.4, are formulated in Section 2. Sections 3 and 4 provide proofs of the theorems, along with some auxiliary facts. The straightforward, but lengthy proof of one of the facts, Proposition 4.2, is presented in Appendix A.

2. Main Results

Our task in this paper is to characterize NEFs with the VF of the form

$$\mathrm{diag}V(m_1,m_2) = \left(Am_1^2 + am_1 + bm_2 + e, Am_2^2 + cm_1 + dm_2 + f\right), \tag{2.1}$$

with $A<0$ and $a,\ldots,f\in\mathbb{R}$. To avoid considering the VFs of the diagonal NEFs analyzed by Bar-Lev et al. in [1], we need to exclude the case $b=c=0$. Due to the symmetry between b and c, in the sequel we shall assume without loss of generality that $b\neq 0$.

We decided to stick here to $A<0$ only and to place our findings on the case $A>0$ in a separate future paper. Our decision is motivated by a rather surprising lack of symmetry between the two cases. It turns out that if $A>0$ then (2.1) implies a much wider and less homogeneous class of parametric bivariate functions that are candidates for the Laplace transforms of the corresponding NEFs, than for $A<0$.

Remark 2.1. An interesting feature of the problem is that it can be formulated equivalently using regression properties. Assume that $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ are independent identically distributed random vectors such that their Laplace transform exists in a set with a nonempty interior. The regression formulation of the problem we solve in this paper is to identify the distributions of $\mathbf{X}$ if

$$\begin{aligned} \mathbb{E}\left((X_1 - Y_1)^2 - 2AX_1Y_1 | \mathbf{X} + \mathbf{Y}\right) &= a(X_1 + Y_1) + b(X_2 + Y_2) + 2e, \\ \mathbb{E}\left((X_2 - Y_2)^2 - 2AX_2Y_2 | \mathbf{X} + \mathbf{Y}\right) &= c(X_1 + Y_1) + d(X_2 + Y_2) + 2f. \end{aligned}$$

For other works in a similar vein, see e.g.[1], [7], [9], [13], [14].

Now we state the main theorem, that specifies the form of a Laplace transform corresponding to the measure generating NEF with the VF given in (2.1).

Theorem 2.2. *Let μ be a probability measure (not concentrated in a point) generating NEF with VF given by* (2.1). *Then the Laplace transform of μ is of the form*

$$L(\theta_1, \theta_2) = \left(\sum_{i=1}^{n_r} \alpha_i \exp\left(\lambda_i \theta_1 + \frac{\lambda_i^2 - a\lambda_i - eA}{b} \theta_2 \right) \right)^{-\frac{1}{A}}, \quad (\theta_1, \theta_2) \in \Theta(\mu), \tag{2.2}$$

where $\{\lambda_1, \dots, \lambda_{n_r}\}$ is the set of distinct real roots of

$$\lambda^4 - 2a\lambda^3 + \left(2Ae + a^2 - db\right)\lambda^2 - \left(2Aae - adb + cb^2\right)\lambda + A^2e^2 - edbA + fb^2A = 0 \tag{2.3}$$

(the coefficients in (2.3) *come from* (2.1)*). Furthermore, $n_r \geq 2$ and either*

- $\alpha_i \geq 0$, $i = 1, \dots, n_r$, $\sum_{i=1}^{n_r} \alpha_i = 1$, $-1/A \in \mathbb{N}$ *and* $\Theta(\mu) = \mathbb{R}^2$, *or*
- $\alpha_i \leq 0$, $i = 1, \dots, n_r$, $\sum_{i=1}^{n_r} \alpha_i = -1$, $-1/A \in 2\mathbb{N}$ *and* $\Theta(\mu) = \mathbb{R}^2$.

Remark 2.3. It is clear that if one of the bullet item conditions above holds then (2.2) is the Laplace transform of a probability measure.

It might be worthy to emphasize that the only assumption we make about $\Theta(\mu)$ is that it contains a neighborhood of the origin (this follows from the fact that μ belongs to NEF). The concrete form of $\Theta(\mu)$ ($\Theta(\mu) = \mathbb{R}^2$) is not assumed but it is implied by the assumptions of Theorem 2.2.

The proof of Theorem 2.2 contains two parts. The first part explains why the distribution sought in (2.1), must have the Laplace transform of the form (2.2). We present this partial result in Proposition 4.2. The second part consists of some arguments concerning the conditions on A, $\alpha_i's$ and $\Theta(\mu)$. The arguments are gathered in a stronger result (see Theorem 2.4 below), which, by a special choice of a matrix Λ in the case of $n_r \geq 3$, implies the conditions on α_i and A (the case $n_r = 2$ will be considered separately in the proof of Theorem 2.2).

The abovementioned stronger result is the following

Theorem 2.4. *Let $r > 0$, $\Lambda = (\Lambda_1, \Lambda_2, \Lambda_3)^T \in \mathbb{R}^{3\times 3}$ and suppose $\Theta \subset \mathbb{R}^3$ contains a neighborhood of the origin. Define*

$$L(\theta) = \left(\alpha_0 + \sum_{i=1}^{3} \alpha_i \exp\langle \Lambda_i, \theta\rangle\right)^r, \quad \theta \in \Theta. \tag{2.4}$$

Assume that equation

$$\mathbf{a}^T \Lambda = \mathbf{0} \tag{2.5}$$

has no solutions in $a = (a_1, a_2, a_3) \in \mathbb{Z}^3$ such that $\exists_{i \neq j}\ a_i a_j < 0$. Then L is Laplace transform of a probability measure if and only if either

(a) $\alpha_i \geq 0$, $i = 0, \dots, 3$, $\sum_{i=0}^{3} \alpha_i = 1$ and $r \in \mathbb{N}$, or
(b) $\alpha_i \leq 0$, $i = 0, \dots, 3$, $\sum_{i=0}^{3} \alpha_i = -1$ and $r \in 2\mathbb{N}$.

Notice that conditions (a) and (b) are sufficient for L to be the Laplace transform of a probability measure.

Remark 2.5. Bar-Lev et al. [Proposition 3.1(a),[1]] consider a similar class of Laplace transforms. If $n \in \mathbb{N}$, define $\mathcal{T}$ to be the family of non-empty subsets of $\{1, 2, \dots, n\}$, and for $\mathbf{z} = (z_1, \dots, z_n)$ put

$$\mathbf{z}^T = \prod_{j \in T} z_j, \quad for \;\; T \in \mathcal{T}.$$

Proposition 3.1 from [1] describes for which α the analytic function of $\mathbf{z}$

$$(1 + \sum_{T \in \mathcal{T}} \alpha_T \mathbf{z}^T)^N, \quad N \in \mathbb{N}, \tag{2.6}$$

has non-negative coefficients in the Taylor expansion in the neighborhood of the origin. For $z_j = e^{\theta_j}$, $j = 1, \dots, n$, this is equivalent to deciding whether (2.6) is the Laplace transform of a positive measure.
Observe that our approach differs from the one taken in [1]. This is because in our problem (2.4), the crucial step is to justify the series expansion of L, while in (2.6) analyticity is assumed. The Λ matrix and the support structure condition (2.5) allow as to identify the masses of atoms among the series coefficients and to infer about the signs of α's. As a consequence of the need to justify the series expansion, we restrict ourselves to $\mathbb{R}^3$. Another difference lies in the fact that the support of the measure corresponding to (2.6) belongs to a lattice in $\mathbb{N}^d$. The support of the measure corresponding to (2.4) depends on the choice of Λ.

Remark 2.6. Again, it might be worthy to emphasize that our assumptions on Θ in Theorem 2.4 are weak. They are dictated by the fact that the measure with Laplace transform (2.4) belongs to a NEF.

3. Proof of Theorem 2.4

Proof of Theorem 2.4. Let us consider first L on $\mathcal{D}$:

$$\mathcal{D} = \left\{ \theta \in \mathbb{R}^3 : \sum_{i=0}^{3} \alpha_i \exp\langle \Lambda_i, \theta \rangle > 0 \right\}.$$

We shall be using the following relation between two vectors $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ in $\mathbb{R}^3$:

$$x \prec y \Longleftrightarrow x_1 < y_1 \vee (x_1 = y_1 \wedge x_2 < y_2) \vee (x_1 = y_1 \wedge x_2 = y_2 \wedge x_3 < y_3).$$

Observe that (2.5) implies the existence of permutation σ of $\{0, 1, 2, 3\}$ such that

$$\Lambda_{\sigma_0} \prec \Lambda_{\sigma_1} \prec \Lambda_{\sigma_2} \prec \Lambda_{\sigma_3}, \tag{3.1}$$

where $\Lambda_0 = (0, 0, 0)$. We will split the reasonings into four cases.

(1) Assume first that $\alpha_{\sigma_0} > 0$. We will show that $\alpha_{\sigma_i} \geq 0$, $i = 1, 2, 3$ and $r \in \mathbb{N}$. Note that (3.1) yields

$$\Lambda_0 \prec \Lambda_{\sigma_1} - \Lambda_{\sigma_0} \prec \Lambda_{\sigma_2} - \Lambda_{\sigma_0} \prec \Lambda_{\sigma_3} - \Lambda_{\sigma_0}. \tag{3.2}$$

This in turn allows us to choose $\theta^\star \in \mathcal{D}$ (negative with arbitrarily large modulus) to ensure that

$$\left| \sum_{i=1}^{3} \frac{\alpha_{\sigma_i}}{\alpha_{\sigma_0}} \exp\langle \Lambda_{\sigma_i} - \Lambda_{\sigma_0}, \theta^\star \rangle \right| < 1.$$

Define $H(\theta) = \exp\langle -r\Lambda_{\sigma_0}, \theta \rangle \frac{L(\theta+\theta^\star)}{L(\theta^\star)}$. ($H$ is the Laplace transform of a probability measure if and only if L is the Laplace transform of a probability measure.). Function H has the following series expansion in the neighborhood of the origin:

$$H(\theta) = C \sum_{j=0}^{\infty} \frac{r(r-1)\cdots(r-j+1)}{j!} \left(\sum_{i=1}^{3} \frac{\alpha_{\sigma_i}}{\alpha_{\sigma_0}} \exp\langle \Lambda_{\sigma_i} - \Lambda_{\sigma_0}, \theta^\star + \theta \rangle \right)^j,$$

where $C = \left(\sum_{i=0}^{3} \frac{\alpha_{\sigma_i}}{\alpha_{\sigma_0}} \exp\langle \Lambda_{\sigma_i} - \Lambda_{\sigma_0}, \theta^\star \rangle \right)^{-r}$.

From (3.2) we see that $Cr\frac{\alpha_{\sigma_1}}{\alpha_{\sigma_0}} \exp\langle \Lambda_{\sigma_1} - \Lambda_{\sigma_0}, \theta^\star \rangle$ is the only coefficient at $\exp\langle \Lambda_{\sigma_1} - \Lambda_{\sigma_0}, \theta \rangle$ in the expansion of H, hence $\alpha_{\sigma_1} \geq 0$. Since (2.5) assures that there is no $a \in \mathbb{N}$, such that $\Lambda_{\sigma_2} - \Lambda_{\sigma_0} = a(\Lambda_{\sigma_1} - \Lambda_{\sigma_0})$, we get that $Cr\frac{\alpha_{\sigma_2}}{\alpha_{\sigma_0}} \exp\langle \Lambda_{\sigma_2} - \Lambda_{\sigma_0}, \theta^\star \rangle$ is the only coefficient at $\exp\langle \Lambda_{\sigma_2} - \Lambda_{\sigma_0}, \theta \rangle$ in the expansion of H, so $\alpha_{\sigma_2} \geq 0$. Analogously, from (2.5) there are no a_1, $a_2 \in \mathbb{N}$ such that $\Lambda_{\sigma_3} - \Lambda_{\sigma_0} = a_1(\Lambda_{\sigma_1} - \Lambda_{\sigma_0}) + a_2(\Lambda_{\sigma_2} - \Lambda_{\sigma_0})$, hence $Cr\frac{\alpha_{\sigma_3}}{\alpha_{\sigma_0}} \exp\langle \Lambda_{\sigma_3} - \Lambda_{\sigma_0}, \theta^\star \rangle$ is the only coefficient at $\exp\langle \Lambda_{\sigma_3} - \Lambda_{\sigma_0}, \theta \rangle$ in the expansion of H, and $\alpha_{\sigma_3} \geq 0$.

Now suppose that $r \notin \mathbb{N}$. Consequently, there exists the smallest k such that $r(r-1)\cdots(r-k+1) < 0$. Since coefficient at

$\exp\langle k(\Lambda_{\sigma_1}-\Lambda_{\sigma_0}),\theta\rangle$ is positive as a mass of an atom in $k(\Lambda_{\sigma_1}-\Lambda_{\sigma_0})$ there exist $n_1,n_2,n_3\in\mathbb{N}$ such that $k(\Lambda_{\sigma_1}-\Lambda_{\sigma_0})=n_1(\Lambda_{\sigma_1}-\Lambda_{\sigma_0})+n_2(\Lambda_{\sigma_2}-\Lambda_{\sigma_0})+n_3(\Lambda_{\sigma_3}-\Lambda_{\sigma_0})$ which contradicts (2.5); hence $r\in\mathbb{N}$.

(2) Now assume that $\alpha_{\sigma_0}\le 0$ and $\alpha_{\sigma_3}>0$. We will show that in fact $\alpha_{\sigma_0}=0$, $\alpha_{\sigma_i}\ge 0$, $i=1,2$ and $r\in\mathbb{N}$.

Note that (3.1) yields

$$\Lambda_{\sigma_0}-\Lambda_{\sigma_3}\prec\Lambda_{\sigma_1}-\Lambda_{\sigma_3}\prec\Lambda_{\sigma_2}-\Lambda_{\sigma_3}\prec\Lambda_0. \tag{3.3}$$

Therefore we can choose $\theta^\star\in\mathcal{D}$ satisfying

$$\left|\sum_{j=0}^{2}\frac{\alpha_{\sigma_j}}{\alpha_{\sigma_3}}\exp\left\langle\Lambda_{\sigma_j}-\Lambda_{\sigma_3},\theta^\star\right\rangle\right|<1$$

and define $G(\theta)=\exp\langle -r\Lambda_3,\theta\rangle\frac{L(\theta+\theta^\star)}{L(\theta^\star)}$. ($G$ is the Laplace transform of a probability measure if and only if L is the Laplace transform of a probability measure.) We can write its series expansion:

$$G(\theta)=C\sum_{j=0}^{\infty}\frac{r(r-1)\cdots(r-j+1)}{j!}\left(\sum_{i=0}^{2}\frac{\alpha_{\sigma_i}}{\alpha_{\sigma_3}}\exp\langle\Lambda_{\sigma_i}-\Lambda_{\sigma_3},\theta+\theta^\star\rangle\right)^j,$$

where $C=\left(\sum_{i=0}^{3}\frac{\alpha_{\sigma_i}}{\alpha_{\sigma_3}}\exp\langle\Lambda_{\sigma_i}-\Lambda_{\sigma_3},\theta^\star\rangle\right)^{-r}$.

From (3.3) we see that $Cr\alpha_{\sigma_3}^{-1}\alpha_{\sigma_2}\exp\langle\Lambda_{\sigma_2}-\Lambda_{\sigma_3},\theta^\star\rangle$ is the only coefficient at $\exp\langle\Lambda_{\sigma_2}-\Lambda_{\sigma_3},\theta\rangle$ in the expansion of G, hence $\alpha_{\sigma_2}\ge 0$. Using the same reasoning as in the preceding case, we conclude from (2.5) that $Cr\alpha_{\sigma_3}^{-1}\alpha_{\sigma_i}\exp\langle\Lambda_{\sigma_i}-\Lambda_{\sigma_3},\theta^\star\rangle$ is the only coefficient at $\exp\langle\Lambda_{\sigma_i}-\Lambda_{\sigma_3},\theta\rangle$ in the expansion of G, hence $\alpha_{\sigma_i}\ge 0$, $i=0,1$, so $\alpha_{\sigma_0}=0$.

In order to conclude that $r\in\mathbb{N}$ it suffices to repeat the previous reasoning.

(3) Now assume that $\alpha_{\sigma_0}\le 0$ and $\alpha_{\sigma_3}<0$ (or $\alpha_{\sigma_0}<0$ and $\alpha_{\sigma_3}\le 0$). We will show that this assumption leads to contradiction.

Since $\alpha_{\sigma_1}+\alpha_{\sigma_2}=1-\alpha_{\sigma_0}-\alpha_{\sigma_3}$, we conclude that $\alpha_{\sigma_1}>\frac{1}{2}$ or $\alpha_{\sigma_2}>\frac{1}{2}$. Without loss of generality we let $\alpha_{\sigma_1}>\frac{1}{2}$.

Again (3.1) implies

$$\Lambda_{\sigma_0}-\Lambda_{\sigma_1}\prec\Lambda_0\prec\Lambda_{\sigma_2}-\Lambda_{\sigma_1}\prec\Lambda_{\sigma_3}-\Lambda_{\sigma_1}.$$

Set $T(\theta)=\exp\langle -r\Lambda_{\sigma_1},\theta\rangle L(\theta)$ and expand it in a neighborhood of the origin:

$$T(\theta)=\alpha_{\sigma_1}^r\sum_{k=0}^{\infty}\frac{r(r-1)\cdots(r-k+1)}{k!\alpha_{\sigma_1}^k}\left(\sum_{i=0,i\ne 1}^{3}\alpha_{\sigma_i}\exp\langle\Lambda_{\sigma_i}-\Lambda_{\sigma_1},\theta\rangle\right)^k.$$

From (2.5), $r\alpha_{\sigma_i}\alpha_{\sigma_1}^{r-1}$ is the only coefficient at $\exp\langle\Lambda_{\sigma_i}-\Lambda_{\sigma_1},\theta\rangle$ in the expansion of T, hence $\alpha_{\sigma_i}\ge 0$ for $i=0,2,3$, so $\alpha_{\sigma_0}=\alpha_{\sigma_3}=0$. It contradicts the assumption.

(4) Now assume that $\alpha_{\sigma_0} = \alpha_{\sigma_3} = 0$. A simplification of the analysis given in cases (1) and (2) leads to $\alpha_{\sigma_1} \geq 0$ and $\alpha_{\sigma_2} \geq 0$. Once we are done with it, the fact that $r \in \mathbb{N}$ follows from the same steps as in the first case.

Summarizing, we have shown that L (defined on D) is a Laplace transform of a probability measure if and only if (a) holds.

Assume now that $r \in 2\mathbb{N}$. If so, we can consider L

$$L(\theta) = \left(\alpha_0 + \sum_{i=1}^{3} \alpha_i \exp\langle \Lambda_i, \theta\rangle\right)^r = \left(-\alpha_0 - \sum_{i=1}^{3} \alpha_i \exp\langle \Lambda_i, \theta\rangle\right)^r,$$

on

$$\mathcal{D}' = \left\{\theta \in \mathbb{R}^3 : \sum_{i=0}^{3} \alpha_i \exp\langle \Lambda_i, \theta\rangle < 0\right\} = \left\{\theta \in \mathbb{R}^3 : \sum_{i=0}^{3} -\alpha_i \exp\langle \Lambda_i, \theta\rangle > 0\right\}.$$

Defining $\tilde{\alpha_i} = -\alpha_i$, $i = 0, \ldots, 3$, from (a) we can conclude that $\tilde{\alpha_i} \geq 0$, hence $\alpha_i \leq 0$, $i = 0, \ldots, 3$. Thus we arrive at (b). $\square$

4. Proof of Theorem 2.2

4.1. **From the form of the diagonal** (2.1) **to the characteristic equation** (2.3)**.** Here we shall justify the transition from the diagonal (2.1) to the characteristic equation (2.3).

Let k be a cumulant function of a measure μ belonging to NEF satisfying (2.1) (that is $k(\theta_1, \theta_2) = \log L(\theta_1, \theta_2)$, $(\theta_1, \theta_2) \in \Theta(\mu)$). Then condition (2.1) can be written equivalently as

$$\frac{\partial^2 k}{\partial \theta_1^2} = A\left(\frac{\partial k}{\partial \theta_1}\right)^2 + a\frac{\partial k}{\partial \theta_1} + b\frac{\partial k}{\partial \theta_2} + e, \tag{4.1}$$

$$\frac{\partial^2 k}{\partial \theta_2^2} = A\left(\frac{\partial k}{\partial \theta_1}\right)^2 + c\frac{\partial k}{\partial \theta_1} + d\frac{\partial k}{\partial \theta_2} + f. \tag{4.2}$$

Define $R = e^{-Ak}$. Then (4.1) and (4.2) become

$$\frac{\partial^2 R}{\partial \theta_1^2} = a\frac{\partial R}{\partial \theta_1} + b\frac{\partial R}{\partial \theta_2} - eAR, \tag{4.3}$$

$$\frac{\partial^2 R}{\partial \theta_2^2} = c\frac{\partial R}{\partial \theta_1} + d\frac{\partial R}{\partial \theta_2} - fAR. \tag{4.4}$$

Since we assume that $b \neq 0$, as a consequence of (4.3) and (4.4) we get

$$\frac{\partial^4 R}{\partial \theta_1^4} - 2a\frac{\partial^3 R}{\partial \theta_1^3} + \left(2Ae + a^2 - db\right)\frac{\partial^2 R}{\partial \theta_1^2} - \left(2aeA - adb + cb^2\right)\frac{\partial R}{\partial \theta_1} + A^2e^2 - edbA + fb^2A = 0, \tag{4.5}$$

with the characteristic equation (2.3).

Remark 4.1. If instead of making the assumption on b, one assumes that $c \neq 0$, then, due to symmetry between b and c, one obtains

$$\frac{\partial^4 R}{\partial \theta_2^4} - 2d\frac{\partial^3 R}{\partial \theta_2^3} + \left(2Af - ac + d^2\right)\frac{\partial^2 R}{\partial \theta_2^2} - \left(2dfA - acd + bc^2\right)\frac{\partial R}{\partial \theta_2} + A^2 f^2 - acfA + ec^2 A = 0, \tag{4.6}$$

which is analogous to (4.5), with the characteristic polynomial

$$\nu^4 - 2d\nu^3 + \left(2Af - ac + d^2\right)\nu^2 - \nu\left(bc^2 - acd + 2dfA\right) + A^2 f^2 - acfA + ec^2 A = 0.$$

The roots of this equation in ν and the roots of (2.3) are connected via

$$\lambda^2 = a\lambda + b\nu - eA, \tag{4.7}$$

$$\nu^2 = c\lambda + d\nu - fA. \tag{4.8}$$

Careful analysis of the roots of (2.3) leads, via various solutions of (4.5), to the following

Proposition 4.2. *Let μ be a probability measure (not concentrated in a point) generating NEF with VF given by* (2.1)*. Let $\{\lambda_1, \ldots, \lambda_{n_r}\}$ be the set of distinct real roots of* (2.3)*. Then $n_r \geq 2$ and the Laplace transform of μ is*

$$L(\theta_1, \theta_2) = \left(\sum_{i=1}^{n_r} \alpha_i \exp\left(\lambda_i \theta_1 + \frac{\lambda_i^2 - a\lambda_i - eA}{b}\theta_2\right)\right)^{-\frac{1}{A}}, \tag{4.9}$$

with $(\theta_1, \theta_2) \in \Theta(\mu)$.

The proof will be presented in Appendix A.1.

Remark 4.3. Analyzing the form of the Laplace transform L given in Proposition 4.2, one can see that the measure corresponding to the case $A = -1$ consists of point masses which are concentrated on a parabola (see (4.7)).

Now, we are in a position to prove our main result.

Proof of Theorem 2.2. Let (X, Y) be a random vector with Laplace transform (2.2). If

$$(\tilde{X}, \tilde{Y}) = (X - \lambda_1/A, bY + aX + eA - \lambda_1^2/A)$$

then the Laplace transform of $(\tilde{X}, \tilde{Y})$ is

$$\tilde{L}(\theta_1, \theta_2) = \left(\alpha_1 + \sum_{i=2}^{n_r} \alpha_i \exp\left[(\lambda_i - \lambda_1)\theta_1 + (\lambda_i^2 - \lambda_1^2)\theta_2\right]\right)^{-\frac{1}{A}}, \tag{4.10}$$

where $(\theta_1, \theta_2) \in \Theta(\tilde{\mu})$ and $\tilde{\mu} = \mathcal{L}(\tilde{X}, \tilde{Y})$. We shall be working with the Laplace transform (4.10) rather than with (2.2) because the form of (4.10) allows us to use Theorem 2.4.

Now we will split our reasonings with respect to the number of distinct roots of (2.3).

First, if $n_r = 3$ or 4, we will indicate how to specify Λ in Theorem 2.4 to make it answer the questions about the coefficients from (4.10) (as a consequence also for (2.2)).

- For $n_r = 4$ we choose

$$\Lambda = \begin{bmatrix} \lambda_2 - \lambda_1 & \lambda_2^2 - \lambda_1^2 & 0 \\ \lambda_3 - \lambda_1 & \lambda_3^2 - \lambda_1^2 & 0 \\ \lambda_4 - \lambda_1 & \lambda_4^2 - \lambda_1^2 & 0 \end{bmatrix}. \tag{4.11}$$

  Obviously this particular Λ plugged into (2.4) yields (4.10) (after changing $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ into $\alpha_1, \alpha_2, \alpha_3, \alpha_4$, respectively). What we need to show is that such Λ satisfies (2.5). To this end we first assume without loss of the generality that $\lambda_2 < \lambda_3 < \lambda_4$. Equation (2.5) can be rewritten equivalently as the system of three equations

$$a_1(\lambda_i - \lambda_1) + a_2(\lambda_i^2 - \lambda_1^2) = 0, \quad i = 2, 3, 4.$$

  From the first two we get

$$a_1 + a_2(\lambda_i + \lambda_1) = 0, \quad i = 2, 3.$$

  Subtracting the first ($i = 1$) from the second one ($i = 2$) yields

$$a_2(\lambda_2 - \lambda_1) = 0,$$

  hence $a_2 = 0$ and $a_1 = 0$. Therefore (4.11) satisfies (2.5).
- For $n_r = 3$ we take

$$\Lambda = \begin{bmatrix} \lambda_2 - \lambda_1 & \lambda_2^2 - \lambda_1^2 & 0 \\ \lambda_3 - \lambda_1 & \lambda_3^2 - \lambda_1^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{4.12}$$

  Analogously to the preceding case, after plugging (4.12) into (2.4) we obtain (4.10) and the same analysis shows that (4.12) satisfies (2.5).

For $n_r = 2$, (4.10) becomes

$$\tilde{L}(\theta_1, \theta_2) = \left(\alpha_1 + \alpha_2 \exp((\lambda_2 - \lambda_1)\theta_1 + (\lambda_2^2 - \lambda_1^2)\theta_2)\right)^{-\frac{1}{A}}.$$

In order to prove the necessary condition, it is enough to analyze one-dimensional Laplace transform l given by

$$l(\theta) = \tilde{L}(\theta/(\lambda_2 - \lambda_1), 0) = (\alpha_1 + \alpha_2 \exp(\theta))^{-\frac{1}{A}}$$

in a neighborhood of 0. Our aim is to show that the necessary conditions for l to be a Laplace transform of a probability measure on $\mathbb{R}$, are either

(1) $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$ and $-\frac{1}{A} \in \mathbb{N}$, or
(2) $\alpha_1, \alpha_2 \leq 0$, $\alpha_1 + \alpha_2 = -1$ and $-\frac{1}{A} \in 2\mathbb{N}$

(sufficiency of the conditions is clear).

Let us consider l on $\mathcal{D}_l = \{\theta : \alpha_1 + \alpha_2 \exp(\theta) > 0\}$, what implies $\alpha_1 + \alpha_2 = 1$ (recall that $l(0) = 1$). Therefore, without loss of generality one can assume that $\alpha_1 \geq 1/2$ (so $\alpha_2 \leq 1/2$). What we want to prove is that $\alpha_1 \geq 1/2$ implies $\alpha_2 \geq 0$. If $\alpha_1 = 1/2$, then $\alpha_2 = 1 - \alpha_1 = 1/2$, so let us assume that $\alpha_1 > 1/2$. In such case $|\alpha_2/\alpha_1| = |(1-\alpha_1)/\alpha_1| < 1$, hence, in a neighborhood of the origin, l can be written as

(4.13)
$$l(\theta) = \alpha_1^{-1/A} \sum_{k=0}^{\infty} \frac{-1/A(-1/A-1)...(-1/A-k+1)}{k!} \left(\frac{\alpha_2}{\alpha_1}\right)^k \exp(k\theta).$$

The only coefficient at $\exp(\theta)$ in (4.13) is $\alpha_2 \alpha_1^{-1/A-1}$, hence $\alpha_2 \alpha_1^{-1/A-1} \geq 0$ and so $\alpha_2 \geq 0$.

Now we focus on the part of conclusion dealing with the exponent $r = -1/A$. We provide here a reasoning analogous to the one given for r in the proof of Theorem 2.4. Suppose that $r \notin \mathbb{N}$. Then there exists the smallest integer k such that $r(r-1)...(r-k+1) < 0$, so the coefficient at $\exp(k\theta)$ in (4.13) is negative. This contradicts the fact that l is a Laplace transform of a discrete probability measure (with non-negative point masses). Thus $r \in \mathbb{N}$. Concluding, we get (1).

In order to get (2), let us now consider l with $r = -1/A \in 2\mathbb{N}$ on $D_l^{'} = \{\theta : \alpha_1 + \alpha_2 \exp(\theta) < 0\}$ (condition $l(\theta) > 0$ is satisfied on $\theta \in D_l^{'}$). Since r is even,

$$l(\theta) = (\alpha_1 + \alpha_2 \exp(\theta))^r = (-\alpha_1 - \alpha_2 \exp(\theta))^r ,$$

and we analyze it on $D_l^{'} = \{\theta : -\alpha_1 - \alpha_2 \exp(\theta) > 0\}$. Denoting $\tilde{\alpha_1} = -\alpha_1$ and $\tilde{\alpha_2} = -\alpha_2$, we arrive at the case considered in (1). Therefore $\tilde{\alpha_1}, \tilde{\alpha_2} \geq 0$, which yields $\alpha_1, \alpha_2 \leq 0$. Thus (2) follows. □

Acknowledgement. I would like to thank Jacek Wesołowski for encouragement and many helpful discussions.

Appendix A.

A.1. **Proof of Proposition 4.2.** In order to prove Proposition 4.2 we will analyze the roots of (2.3) and examine corresponding R functions. Our aim will be to eliminate the solutions of (4.5) that do not lead (via $L = R^{-1/A}$) to Laplace transforms of probability measures. To do so, we will treat R as a function of θ_1 only (with fixed θ_2), and we shall use auxiliary Lemma A.1 and Lemma A.2 presented below to reject some inadmissible solutions.

Lemma A.1. *Let $r > 0$. Assume that P_m is a polynomial of degree m over $\mathbb{R}$ and define*

(A.1)
$$f(\theta) = P_m(\theta) + \sum_{i=1}^{k} A_i \exp(\lambda_i \theta) + B\theta \exp(\gamma\theta), \quad \theta \in \Theta,$$

where Θ contains some neighborhood of zero, $\lambda_1 < \lambda_2 < \ldots < \lambda_k$ (and none of them is zero), and γ, B, A_i are some real numbers.
If f^r is a Laplace transform of a probability measure then $P_m \equiv P_0$ and $B = 0$.

Proof. Since f is defined in a neighborhood of zero, we consider a characteristic function $\phi(t) = f^r(it)$. We have

$$|\phi(t)|^2 = \left| P_m(it) + Bit\exp(\gamma it) + \sum_{i=1}^{k} A_i \exp(\lambda_i it) \right|^{2r}.$$

Function $|\phi|$ is bounded on $\mathbb{R}$ as the absolute value of a characteristic function of a probability measure. Therefore $P_m \equiv P_0$ and $B = 0$. $\square$

Lemma A.2. *Let $r > 0$ and*

$$f(\theta) = \sum_{j=0}^{1} e^{\lambda_j\theta} \left[(A_{0j} + \theta B_{0j})\cos(\gamma_j\theta) + (A_{1j} + \theta B_{1j})\sin(\gamma_j\theta) \right] \\ + \sum_{j=2}^{3} A_j \exp(\lambda_j\theta), \quad \theta \in \Theta,$$

where Θ contains some neighborhood of zero, be a function with all the parameters being some real numbers. Furthermore, assume that $\lambda_0 + i\gamma_0 \neq \lambda_1 + i\gamma_1$. Then if f^r is a Laplace transform of a probability measure, then $A_{0j} = A_{1j} = B_{0j} = B_{1j} = 0$, $j = 0, 1$.

Proof. Let $\phi(t) = f^r(it)$ be a characteristic function corresponding to f^r, then

$$|\phi(t)|^2 = \\ \left| \sum_{j=0}^{1} \left(e^{\lambda_j it}(A_{0j} + itB_{0j}) \frac{e^{\gamma_j t} + e^{-\gamma_j t}}{2} + e^{\lambda_j it}(A_{1j} + itB_{1j}) \frac{e^{-\gamma_j t} - e^{\gamma_j t}}{2i} \right) \right. \\ \left. + A_2 \exp(i\lambda_2 t) + A_3 \exp(i\lambda_3 t) \right|^{2r}.$$

Since $|\phi|$ is bounded on $\mathbb{R}$, we arrive at the conclusion. $\square$

Proof of Proposition 4.2. We shall separately consider all situations regarding the roots of (2.3).

A.1.1. *Four single real roots of* (2.3). Let $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ be the roots of (2.3). Then R is of the form

$$R(\theta_1, \theta_2) = \sum_{i=1}^{4} A_i(\theta_2) \exp(\lambda_i\theta_1), \tag{A.2}$$

where $A_i(\cdot)$, $i = 0, \dots, 3$ are some real functions. Plugging (A.2) into (4.4) we obtain

$$\lambda_i^2 A_i(\theta_2) - a\lambda_i A_i(\theta_2) - bA_i^{'}(\theta_2) - eAA_i(\theta_2) = 0, \quad i = 1, 2, 3, 4.$$

Therefore the explicit formulas for $A_i(\cdot)$'s are

$$A_i(\theta_2) = A_i \exp\left(\frac{\lambda_i^2 - a\lambda_i - eA}{b}\theta_2\right), \ i = 1, 2, 3, 4,$$

and A_i, $i = 1, 2, 3, 4$ are some real constants. Hence

$$R(\theta_1, \theta_2) = \sum_{i=1}^{4} A_i \exp\left(\lambda_i \theta_1 + \frac{\lambda_i^2 - a\lambda_i - eA}{b}\theta_2\right). \tag{A.3}$$

A.1.2. *One double and two single real roots of* (2.3). Let λ_1, λ_2 be single roots and λ_3 a double root of (2.3). Then R is of the form

$$R(\theta_1, \theta_2) = \sum_{i=1}^{2} A_i(\theta_2) \exp(\lambda_i \theta_1) + (A_3(\theta_2) + A_4(\theta_2)\theta_1) \exp(\lambda_3 \theta_1), \tag{A.4}$$

where $A_i(\cdot)$, $i = 1, 2, 3, 4$, are some real functions. Since $R^{-1/A}$ is a Laplace transform of a probability measure, using Lemma A.1 we conclude that $A_4 \equiv 0$. Plugging (A.4) into (4.3) we obtain

$$\lambda_i^2 A_i(\theta_2) - a\lambda_i A_i(\theta) - bA_i^{'}(\theta_2) - eAA_i(\theta_2) = 0, \ i = 1, 2, 3.$$

These yield

$$A_i(\theta_2) = A_i \exp\left(\frac{\lambda_i^2 - a\lambda_i - eA}{b}\theta_2\right), \ i = 1, 2, 3,$$

and

$$R(\theta_1, \theta_2) = \sum_{i=1}^{3} A_i \exp\left(\lambda_i \theta_1 + \frac{\lambda_i^2 - a\lambda_i - eA}{b}\theta_2\right).$$

A.1.3. *One single and one triple real roots of* (2.3). Let λ_1 be a single and λ_2 a triple real root of (2.3). Then

$$R(\theta_1, \theta_2) = A_1(\theta_2) \exp(\lambda_1 \theta_1) + \left(A_2(\theta_2) + A_3(\theta_2)\theta_1 + A_4(\theta_2)\theta_1^2\right) \exp(\lambda_2 \theta_1),$$

where $A_i(\cdot)$, $i = 1, 2, 3, 4$, are some real functions. Analogously as in the previous subsection, using Lemma A.1 (for $R(\theta_1, \theta_2)\exp(-\lambda_2\theta_1)$), we conclude that $A_3 \equiv A_4 \equiv 0$. Function R in such case is of the form

$$R(\theta_1, \theta_2) = \sum_{i=1}^{2} A_i \exp\left(\lambda_i \theta_1 + \frac{\lambda_i^2 - a\lambda_i - eA}{b}\theta_2\right),$$

where A_i, $i = 1, 2$ are some real functions.

A.1.4. *Quadruple real root of* (2.3). Let λ_1 be a quadruple real root of (2.3). Then

$$R(\theta_1, \theta_2) = A_1 \exp\left(\lambda_1\theta_1 + \frac{\lambda_1^2 - a\lambda_1 - eA}{b}\theta_2\right),$$

where A_1 is a real constant.

A.1.5. *Two double real roots of* (2.3). Let λ_1 and λ_2 be two distinct double roots of (2.3). Steps analogous to the ones taken in Section A.1.2, with the help of Lemma A.1, yield

$$R(\theta_1, \theta_2) = \sum_{i=1}^{2} A_i \exp\left(\lambda_i\theta_1 + \frac{\lambda_i^2 - a\lambda_i - eA}{b}\theta_2\right),$$

where A_i, $i = 1, 2$, are some real constants.

A.1.6. *Two single real and two single complex roots of* (2.3). Let λ_1, λ_2 be two real and $\lambda_3 + i\gamma_3$, $\lambda_3 - i\gamma_3$ be two complex roots of (2.3). Then R takes the form

$$R(\theta_1, \theta_2) = \sum_{i=1}^{2} A_i(\theta_2) \exp\left(\lambda_i\theta_1 + \frac{\lambda_i^2 - a\lambda_i - eA}{b}\theta_2\right) + \\ \exp(\lambda_3\theta_1)\left(A_3(\theta_2)\cos(\gamma_3\theta_1) + A_4(\theta_2)\sin(\gamma_3\theta_1)\right).$$

From Lemma A.2 we conclude that $A_3 \equiv A_4 \equiv 0$, hence

$$R(\theta_1, \theta_2) = \sum_{i=1}^{2} A_i \exp\left(\lambda_i\theta_1 + \frac{\lambda_i^2 - a\lambda_i - eA}{b}\theta_2\right),$$

where A_i, $i = 2, 3$ are some real constants.

A.1.7. *Four distinct complex roots of* (2.3). Let $\lambda_1 + i\gamma_1$, $\lambda_1 - i\gamma_1$, $\lambda_2 + i\gamma_2$ and $\lambda_2 - i\gamma_2$ be the roots of (2.3). Then R takes the form

$$R(\theta_1, \theta_2) = \exp(\lambda_1\theta_1)\left(A_1(\theta_2)\cos(\gamma_1\theta_1) + A_2(\theta_2)\sin(\gamma_1\theta_1)\right) + \\ \exp(\lambda_2\theta_1)\left(A_3(\theta_2)\cos(\gamma_2\theta_1) + A_4(\theta_2)\sin(\gamma_2\theta_1)\right).$$

Using Lemma A.2 we conclude that there does not exist probability measure with the Laplace transform $L = R^{-1/A}$.

A.1.8. *Two double complex roots of* (2.3). Analogously to the preceding case, there is no probability measure corresponding to R

$$R(\theta_1, \theta_2) = \exp(\lambda_1\theta_1)\Big[\cos(\gamma_1\theta_1)\big(A_1(\theta_2) + \theta_1 A_2(\theta_2)\big) + \\ \sin(\gamma_1\theta_1)\big(A_3(\theta_2) + \theta_1 A_4(\theta_2)\big)\Big],$$

where $\lambda_1 + i\gamma_1$ and $\lambda_1 - i\gamma_1$ are double complex roots of (2.3) and A_i, $i = 1, 2, 3, 4$, are some real functions.

A.1.9. *Double real root and two single complex roots of* (2.3). Let λ_1, $\lambda_2+i\gamma_2$ and $\lambda_2-i\gamma_2$ be the roots of (2.3), then

$$R(\theta_1,\theta_2)=\exp(\lambda_1\theta_1)\left(A_1(\theta_2)+\theta_1A_2(\theta_2)\right) \\ +\exp(\lambda_2\theta_1)\left(A_3(\theta_2)\cos(\gamma_2\theta_1)+A_4(\theta_2)\sin(\gamma_2\theta_1)\right).$$

From Lemma A.1 and Lemma A.2 we obtain $A_2\equiv A_3\equiv A_4\equiv 0$. Hence R takes the form

$$R(\theta_1,\theta_2)=\exp\left(\lambda_1\theta_1+\frac{\lambda_1^2-a\lambda_1-eA}{b}\theta_2\right).$$

The conclusion of Proposition 4.2 follows by a straightforward aggregation of the above points. □

References

[1] S.K. Bar-Lev, D. Bshouty, P. Enis, G. Letac, Lu I-Li and D. Richards, *The diagonal multivariate natural exponential families and their classification,* J. Theoret. Probab. 7 (1994), pp. 883–929.

[2] S. K. Bar-Lev, *Methods of constructing characterizations by constancy of regression on the sample mean and related problems for NEF's*, Math. Methods Statist. 16 (2) (2007), pp. 96-109.

[3] P. Bernardoff, *Which negative multinomial distributions are inifinitely divisible*, Bernoulli 9(5) (2003), pp. 877-893.

[4] M. Casalis, *Les familles exponentielles sur R^2 de fonction-variance $V(m) = am\otimes m+B(m)+C$*, C. R. Acad. Sci. Paris 314 (1992), pp. 635-638.

[5] M. Casalis, *The 2d + 4 simple quadratic natural exponential families on R^d*, Ann. Statist. 24 (1996), pp. 1828-1854.

[6] J.Chachulska, *Bivariate natural exponential families with linear diagonal variance functions*, PMS 30 (1) (2010), pp.121-139.

[7] E. B. Fosam and D. N. Shanbhag, *An extended Laha-Lukacs characterization result based on a regression property*,

[8] A.Hassairi and M.Zarai, *Characterization of the simple cubic multivariate exponential families*, J.Funct.Anal.235(2006), pp.69-89.

[9] F. S. Gordon, *Characterizations of populations using regression properties*, Ann. Statist. 1 (1973), pp. 114-126.

[10] C. Kokonendji, V. Seshadri *On the determinant of the second derivative of a Laplace transform*, Ann. Statist. 24(1998), pp. 1813-1827.

[11] C. Kokonendji and A. Masmoudi, *A characterization of Poisson-Gaussian families by generalized variance*, Bernoulli 12 (2) (2006), pp. 371-379.

[12] T.J. Kozubowski, K. Podgórski, *Rational characteristic functions and geometric infinite divisibility*, J. Math. Anal. Appl. 365 (2010), no. 2, 625?-637.

[13] R. G. Laha, E. Lukacs, *On a problem connected with quadratic regression*, Biometrika 47 (1960), pp. 335–343.

[14] G. Letac, J. Wesołowski, *Laplace transforms which are negative powers of quadratic polynomials,* T. Am. Math. Soc. 360 (2008), pp. 6475-6496.

[15] G. Letac, *Lectures on natural exponential families and their variance functions*, Monograph No 50, IMPA, Rio de Janeiro 1992.

[16] G. Letac, *Le problème de la classification des familles exponentielles naturelles de $\mathbf{R^d}$ ayant une fonction-variance quadratique*, in: Probability Measures on Groups IX, Lecture Notes in Math. No 1379 (1989).

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska, Koszykowa 75, 00-662 Warsaw, Poland

E-mail address: J.Matysiak@mini.pw.edu.pl